Bayesian model selection for electromagnetic kaon production on the nucleon

L. De Cruz[*,a], D. G. Ireland[b], P. Vancraeyveld[a], J. Ryckebusch[a]

[a]*Department of Physics and Astronomy, Ghent University, Proeftuinstraat 86, B-9000 Gent, Belgium*
[b]*Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

Abstract

We present the results of a Bayesian analysis of a Regge model to describe the background contribution for $K^+\Lambda$ and $K^+\Sigma^0$ photoproduction. The model is based on the exchange of $K^+(494)$ and $K^{*+}(892)$ trajectories in the t-channel. We utilise the Bayesian evidence $\mathcal{Z}$ to determine the best model variant for each channel. The Bayesian evidence integrals were calculated using the Nested Sampling algorithm. For different prior widths, we find decisive Bayesian evidence ($\Delta \ln \mathcal{Z} \approx 24$) for a $K^+\Lambda$ photoproduction Regge model with a positive vector coupling and a negative tensor coupling constant for the $K^{*+}(892)$ trajectory, and a rotating phase factor for both trajectories. Using the χ^2 minimisation method, one could not draw this conclusion from the same dataset. For the $K^+\Sigma^0$ photoproduction Regge model, on the other hand, the difference between the evidence integrals is insufficient to pinpoint one model variant.

Key words: Regge phenomenology, Bayesian inference, Model selection
PACS: 11.55.Jy, 12.40.Nn, 13.60.Le

1. Introduction

To resolve the structure of the nucleon, many models based on effective degrees of freedom have been developed. One of the more popular ones is the constituent quark model (CQM), which presents the nucleon as a system of three constituent quarks [1]. The CQM, however, predicts far more resonances than confirmed by experiment. This may lead one to turn to other models that predict fewer resonances [2].

Alternatively, this discrepancy may arise because missing resonances do not couple to the channels commonly used for nucleon spectroscopy, such as the pion-nucleon (πN) channel. This issue can be addressed by using an electromagnetic probe instead of a pion, and by examining decay channels other than

---

πN. Hence, electromagnetic open strangeness or kaon-hyperon (KY) production is suggested as a key process to seek for unobserved resonances. The analysis of this process faces other difficulties, such as a small cross section and a high threshold [3, 4]. Unlike threshold one-pion photoproduction, which is dominated by the $\Delta(1232)$ resonance, the KY reaction channel opens in a resonance-rich energy region. The identification of those resonances constitutes a major challenge in modelling KY production. A second characteristic of this channel is the great importance of the non-resonant or background contributions. Hence, a correct determination of the background is crucial for a correct assessment of the resonance contributions.

The earliest studies of the KY production focused on the estimation of coupling constants within a single model variant [5, 6]. Only when the emphasis came to lie on identifying missing resonances did the focus shift from parameter estimation to model comparison. The statistical tools, however, have not been adapted to this new objective. The least-squares method in particular has often been stretched beyond its limits, being used not only as an optimisation tool, but also as a model selection criterion. In this Letter, we present the Bayesian evidence computation, a method based on the principles of Bayesian inference, as a more robust and well-founded tool for model comparison.

The outline of this Letter is as follows. The next section introduces the Bayesian evidence and the Nested Sampling (NS) algorithm for evidence computation. The effectiveness of this algorithm is subsequently demonstrated for a Regge model in Section 3 and the results are discussed in Section 4. The conclusions and outlook are given in Section 5.

2. Bayesian analysis

Bayesian analysis is an established tool for model selection in astronomy and cosmology, and is gaining momentum in other fields [7, 8]. The potential of this method in hadronic physics was recently demonstrated in a Bayesian analysis of pentaquark data by Ireland et al. [9], and for parameter estimation in effective field theories by Schindler et al. [10]. The quantity of interest for model comparison is the Bayesian evidence, which will be derived below.

One can straightforwardly derive the posterior probability of a model M, given a set of experimental data $\{d_k\}$. Indeed, using Bayes' theorem, one can write this probability as

$$P(M|\{d_k\}) = \frac{P(\{d_k\}|M)\,P(M)}{P(\{d_k\})}. \tag{1}$$

The quantity $P(\{d_k\}|M)$ is referred to as the marginal likelihood or the Bayesian evidence ($\mathcal{Z}$). If the model M can be parametrised with a set $\boldsymbol{\alpha_M}$, this probability can be written as an integral over all possible values of these parameters. This procedure, which is referred to as marginalisation, yields the following

Table 1: Jeffreys' scale for the natural logarithms of evidence ratios $\Delta \ln \mathcal{Z} = \ln \frac{\mathcal{Z}_\mathcal{A}}{\mathcal{Z}_\mathcal{B}}$ [11, 12].

	$|\Delta \ln \mathcal{Z}|$	< 1	Not worth more than a bare mention
$1 <$	$|\Delta \ln \mathcal{Z}|$	< 2.5	Significant
$2.5 <$	$|\Delta \ln \mathcal{Z}|$	< 5	Strong to very strong
$5 <$	$|\Delta \ln \mathcal{Z}|$		Decisive

expression for the Bayesian evidence:

$$\mathcal{Z} \equiv P(\{d_k\}\,|M) \tag{2}$$

$$= \int P(\{d_k\}\,, \boldsymbol{\alpha_M}|M)\, d\boldsymbol{\alpha_M} \tag{3}$$

$$= \int \underbrace{P(\{d_k\}\,|\boldsymbol{\alpha_M}, M)}_{(i)}\, \underbrace{P(\boldsymbol{\alpha_M}|M)}_{(ii)}\, d\boldsymbol{\alpha_M}. \tag{4}$$

Eq. (4) states that the Bayesian evidence is the integral of the product of two distributions: (i) the probability of the dataset $\{d_k\}$, given the set of parameters $\boldsymbol{\alpha_M}$ and the model M, and (ii) the probability of the set of parameters $\boldsymbol{\alpha_M}$, given the model M. The first factor, $P(\{d_k\}\,|\boldsymbol{\alpha_M}, M)$, can be identified as the likelihood function, $\mathcal{L}(\boldsymbol{\alpha_M})$. Any prior knowledge of the parameters' probability distribution before considering the data $\{d_k\}$ is contained in the second factor $P(\boldsymbol{\alpha_M}|M)$. This distribution, which is indispensable in Bayesian statistics, is referred to as the prior distribution $\pi(\boldsymbol{\alpha_M})$. These two substitutions allow us to write the evidence in a more familiar form,

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{\alpha_M})\, \pi(\boldsymbol{\alpha_M})\, d\boldsymbol{\alpha_M}, \tag{5}$$

in which the explicit dependence on $\{d_k\}$ and M is omitted for brevity.

It is clear that the actual quantity of interest for model comparison is the relative probability of a model M_A versus a model M_B, given the available experimental data $\{d_k\}$. Writing down the probability ratios and subsequently applying Bayes' theorem, one can see how the evidence emerges from this expression:

$$\frac{P(M_A|\,\{d_k\})}{P(M_B|\,\{d_k\})} = \frac{P(\{d_k\}\,|M_A)}{P(\{d_k\}\,|M_B)}\, \frac{P(M_A)}{P(M_B)} \tag{6}$$

$$= \frac{\mathcal{Z}_\mathcal{A}}{\mathcal{Z}_\mathcal{B}} \quad \text{for } P(M_A) = P(M_B). \tag{7}$$

Any prior preference for one model over the other can be incorporated by the factor $\frac{P(M_A)}{P(M_B)}$. As we have no prior preference for one of the models, we can take this value to be one, hence reducing the comparison of two models to the calculation of the evidence ratio, which is often referred to as the Bayes

factor. The direct relation between $\mathcal{Z}$ and a model's probability elucidates the term "evidence": if a model has a higher value of $\mathcal{Z}$, there is more evidence in favour of this model. In accordance with our intuitive notions, evidence is not only based on experimental data, but also on theoretical restrictions that are incorporated through the prior distribution. The natural logarithm of the evidence ratio can be interpreted qualitatively with the aid of Jeffreys' scale, listed in Table 1.

The analytical form of the likelihood function $\mathcal{L}(\boldsymbol{\alpha_M})$ is rarely known and a normal distribution is often used to approximate it. Indeed, data points are independent and are usually reported to have normally distributed errors. This gives rise to a χ^2-distribution for the quantity defined as

$$\chi^2(\boldsymbol{\alpha_M}) = \sum_i^N \frac{(d_i - f_i(\boldsymbol{\alpha_M}))^2}{2\sigma_i^2}, \tag{8}$$

where σ_i is the error bar of data point d_i, and $f_i(\boldsymbol{\alpha_M})$ is the corresponding model prediction. The χ^2-distribution can be approximated by a normal distribution if the number of degrees of freedom k is sufficiently large. This value is defined as $k = N - \dim(\boldsymbol{\alpha_M})$, which approximates the number of data points for a sufficiently large dataset and low number of free parameters. The expression for this limiting distribution is [13]:

$$\mathcal{L}(\boldsymbol{\alpha_M}) \approx \frac{1}{2\sqrt{\pi k}} \exp{-\frac{(\chi^2(\boldsymbol{\alpha_M}) - k)^2}{4k}}. \tag{9}$$

In an analysis based on χ^2 minimisation – which is an approximation to a maximum likelihood fit – only the maximum value of $\mathcal{L}(\boldsymbol{\alpha_M})$ is considered for model selection. A Bayesian approach is more comprehensive, as it evaluates the model over its entire parameter space, and takes the prior distribution into account. This distinction is illustrated in Fig. 1.

Determining the evidence of a model is not a straightforward task, because it requires the calculation of multidimensional integrals of the type (5). Most often, analytical simplifications are not possible and it is key to adopt numerical integration techniques that are optimised for the problem at hand.

Nested Sampling (NS) is a novel integration technique for computing Bayesian evidence, developed by Skilling [14, 15]. This technique significantly reduces the computational cost of the integral over the model's parameter space by transforming it into a one-dimensional integral over the prior mass $dX = \pi(\boldsymbol{\alpha_M})d\boldsymbol{\alpha_M}$. This is accomplished by regarding the prior mass as a monotonically decreasing function of the likelihood, λ:

$$X(\lambda) = \int_{\mathcal{L}(\boldsymbol{\alpha_M}) > \lambda} \pi(\boldsymbol{\alpha_M})d\boldsymbol{\alpha_M}. \tag{10}$$

Assuming a normalised prior, we can hence write the evidence as the following integral over $\mathcal{L}(X)$, the inverse of $X(\lambda)$,

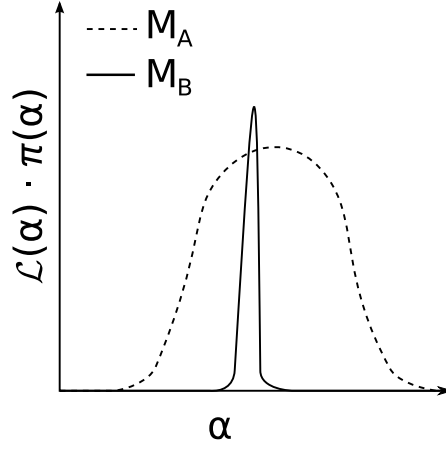$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)\,dX, \tag{11}$$

Figure 1: $\mathcal{L}(\alpha) \cdot \pi(\alpha)$ for two different single-parameter models M_A and M_B. The traditional least-squares method would favour M_B, as it only takes into account the maximum value of the likelihood. In contrast, a Bayesian approach would favour model M_A, as its evidence, given by Eq. (5), is greater.
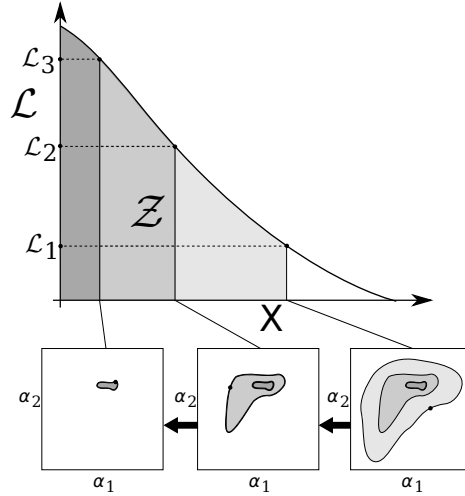

Figure 2: The Bayesian evidence $\mathcal{Z}$ as an integral of the likelihood $\mathcal{L}$ over the prior mass X. The lower plots illustrate that for a uniform prior, the increasing prior masses $X(\mathcal{L}_i)$ are equal to the area or mass $\{\,\boldsymbol{\alpha}\,|\,\mathcal{L}(\boldsymbol{\alpha}) > \mathcal{L}_i\,\}$ inside nested iso-likelihood contours. The thick arrows indicate the order of integration.

as illustrated in Fig. 2. This integral can be approximated by a sum of likelihoods, weighted with their respective prior mass contribution ΔX. In NS, this sum is computed using Markov chain Monte Carlo methods to sample from the parameter space with the constraint $\mathcal{L}_{i+1} > \mathcal{L}_i$. The prior mass of a sample of N points with likelihood greater than $\mathcal{L}_i$ can be estimated by a factor $\exp -\frac{i}{N}$. This can be derived from the probability distribution of the largest value t_N of a set of N uniformly distributed points in the interval $[0, 1]$ (representing the total prior weight):

$$P_N(t) = N t^{N-1}. \tag{12}$$

Indeed, the estimation value for $\langle \log t \rangle_N$ is $-\frac{1}{N}$, which after i iterations amounts to a sum of $-\frac{i}{N}$.

Besides efficiently calculating the evidence, NS can also be used for determining the posterior distribution and, more importantly, for parameter estimation. A more detailed account of the technique as well as implementation examples can be found in Skilling's work [15].

In the following section, we will present the results of an application of the NS technique to a Regge model for $K^+\Lambda$ and $K^+\Sigma^0$ production.

3. Bayesian analysis of a Regge Model

Regge phenomenology is a powerful tool to economically describe reactions at high energies [16]. A Regge model based on the exchange of charged meson Regge trajectories was applied to KY photoproduction by Vanderhaeghen, Guidal and Laget [17]. They found the exchange of two t-channel trajectories sufficient to successfully describe the cross sections as well as polarisation observables in photoproduction of kaons above the resonance region [18].

In the Regge-plus-resonance (RPR) description of electromagnetic KY production, developed by Corthals et al., the Regge background is complemented with s-channel nucleon (N^*) resonances. This hybrid approach ensures a correct high-energy behaviour as well as an improved description of the resonance region [4, 19, 20].

In the case of $K^+\Lambda$ and $K^+\Sigma^0$ photoproduction, the Regge background can be modelled with the exchange of the $K^+(494)$ and $K^{*+}(892)$ trajectories. This amplitude is derived from the t-channel Feynman amplitude by replacing the Feynman propagator by the respective Regge propagator [4]:

$$
\begin{aligned}
\mathcal{P}_{Regge}^{K^+}(s,t) = {} & \left(\frac{s}{s_0}\right)^{\alpha_K(t)} \frac{1}{\sin\big(\pi\alpha_K(t)\big)} \\
& \times \frac{\pi\alpha_K'}{\Gamma\big(1+\alpha_K(t)\big)} \left\{ \begin{matrix} 1 \\ e^{-i\pi\alpha_K(t)} \end{matrix} \right\}
\end{aligned}
\tag{13}
$$

6

$$\mathcal{P}_{Regge}^{K^{*+}}(s,t) = \left(\frac{s}{s_0}\right)^{\alpha_{K^*}(t)-1} \frac{1}{\sin\left(\pi\alpha_{K^*}(t)\right)}$$

$$\times \frac{\pi\alpha'_{K^*}}{\Gamma\left(\alpha_{K^*}(t)\right)} \left\{ \begin{array}{c} 1 \\ e^{-i\pi\alpha_{K^*}(t)} \end{array} \right\}. \tag{14}$$

The kaon trajectories are given by [4]:

$$\alpha_K(t) = 0.70 \text{ GeV}^{-2}\left(t - m_K^2\right) \tag{15}$$

$$\alpha_K^*(t) = 1 + 0.85 \text{ GeV}^{-2}\left(t - m_{K^*}^2\right), \tag{16}$$

and the scale factor $s_0 = 1 \text{ GeV}^{-2}$. The so-called sign factor in the Regge propagator is reduced to a phase factor of either 1 (constant phase) or $e^{-i\pi\alpha(t)}$ (rotating phase) due to the strong degeneracy of the trajectories. This assumption is inspired by the structureless high-energy differential cross-section. These phases cannot be determined on theoretical grounds. The possibility of the K^+ and K^{*+} trajectories having a constant phase is excluded, as this combination gives rise to a photon asymmetry $\Sigma = 0$, which disagrees with the data. The remaining three possibilities, namely rotating K^+ /rotating K^{*+}, rotating K^+/constant K^{*+}, and constant K^+/rotating K^{*+}, will be abbreviated as rot./rot., rot./cst. and cst./rot., respectively.

Apart from the three choices with regard to the phases, the model has three continuous parameters. These are the strong coupling constant g_{K^+Yp} of the K^+ trajectory and the tensor and vector couplings of the K^{*+} trajectory,

$$G_{K^{*+}}^{v,t} = \frac{e\, g_{K^{*+}Yp}^{v,t}}{4\pi}\, \kappa_{K^+K^{*+}}\,. \tag{17}$$

Here, $\kappa_{K^+K^{*+}}$ is the transition magnetic moment for $K^{*+} \to \gamma K^+$ decay.

In the following analysis, the Regge background will be constrained by the high-energy data, in accordance with the RPR-approach. In this energy region, there is a set of 72 data points for the $K^+\Lambda$ channel, comprising 56 differential cross section data points ($\frac{d\sigma}{dt}$) [21], 9 photon asymmetries (Σ) [22] and 7 recoil asymmetries (P) [23]. The database for $K^+\Sigma^0$ photoproduction at high energies is even smaller, with only 48 differential cross section data points [21] and 9 photon asymmetries (Σ) [22]. Optimisation of the above-mentioned parameters against these data reveals that there are several model variants with comparable χ^2 values [4]. For example, in both $K^+\Lambda$ and $K^+\Sigma^0$ photoproduction, the signs of $G_{K^{*+}}^v$ and $G_{K^{*+}}^t$ cannot be established conclusively using the χ^2-method [4].

The sign and phase ambiguities may not seem important for the Regge model itself. For the RPR model, however, an exact determination of the background parameters is of major importance, because it affects the extraction of the resonance information.

Ref. [20] shows that by comparing the RPR model variants for electromagnetic $K^+\Lambda$ production to photo- and electroproduction data from the resonance

region, all but one Regge background model can be eliminated. We will show that the Bayesian evidence can be used to distinguish among the twelve different models that result from the possible sign and phase combinations, using only the high-energy dataset.

4. Results

The prior $\pi(\boldsymbol{\alpha})$ is chosen to be a uniform distribution. Note that under conditions of highly concentrated likelihood, for which the prior distribution varies mildly, the likelihood dominates the shape of the posterior distribution [14]. Accordingly, the evidence calculations will not be largely affected by the choice with regard to the prior distribution. This means that a uniform distribution will lead to results that are similar to those obtained with a Gaussian or any other well-behaved distribution. We show that the bulk of the likelihood is indeed concentrated at parameter values below 100 by demonstrating that evidence calculations for prior widths equal to 100 and much greater than 100 yield the same results.

Prior information exists for the coupling constants of the K^+Y^0p vertices. Indeed, the following relations follow from SU(3) symmetry [24]:

$$g_{K^+\Lambda p} = -\frac{1}{\sqrt{3}}(3 - 2\alpha)\, g_{\pi NN} \tag{18}$$

$$g_{K^+\Sigma^0 p} = (2\alpha - 1)\, g_{\pi NN} \tag{19}$$

where $\alpha = F/(F + D)$ quantifies the ratio of F-type to D-type coupling and $g_{\pi NN}$ is the pion-nucleon coupling constant. It is commonly assumed that SU(3) symmetry can be broken at the 20% level [25]. Inserting the experimentally determined values $\alpha = 0.644$ and $g_{\pi NN}^2/\sqrt{4\pi} = 14.3$ yields the following prior ranges for the coupling constants of the K^+Y^0p vertices [24, 25]:

$$-4.5 \leq \frac{g_{K^+\Lambda p}}{\sqrt{4\pi}} \leq -3.0 \tag{20}$$

$$0.9 \leq \frac{g_{K^+\Sigma^0 p}}{\sqrt{4\pi}} \leq 1.3. \tag{21}$$

There are no reliable theoretical constraints for the $K^{*+}Y^0p$ vertices [26]. We therefore choose a uniform distribution between zero and a value much larger than the natural value of one. To justify this choice for the prior interval, a sensitivity analysis of the evidence ratios is performed by repeating the calculations for different prior ranges.

The results of these calculations for the $K^+\Lambda$ and $K^+\Sigma^0$ production models are displayed in Table 2 and 3 respectively. These tables list the computed values of $\Delta \ln \mathcal{Z} = \ln \mathcal{Z} - \ln \mathcal{Z}_{max}$, using a prior width of respectively 100, 1000 and 10000. Changing the prior width from 100 to 1000 results in a difference of less than 5% in the computed values. For a prior width of 10000, the error increases significantly due to a reduced sampling efficiency. More importantly,

however, the ranking of the models is not significantly affected. Clearly, the effect of the prior width on the relative probabilities of the models is negligible, provided that it is large enough to contain the area where most of the likelihood is concentrated.

Table 2: Logarithms of the evidence ratios ($\Delta \ln \mathcal{Z} \equiv \ln (\mathcal{Z}/\mathcal{Z}_{max})$) for the twelve model variants resulting from phase and sign ambiguities in the two-trajectory Regge model for $K^+\Lambda$ photoproduction. The results are listed in order of decreasing probability for a prior width of 100.

G^v	G^t	K^+/K^{*+} phase	$\pi = U(0, \pm 100)$			$\pi = U(0, \pm 1000)$			$\pi = U(0, \pm 10000)$		
+	−	rot. / rot.	0			0			0		
−	−	rot. / cst.	−24.30	±	0.75	−24.7	±	3.5	−24	±	41
+	+	rot. / rot.	−77.23	±	0.76	−77.3	±	2.9	−79	±	59
−	+	rot. / cst.	−387.04	±	0.77	−387.7	±	4.2	−411	±	68
+	+	rot. / cst.	−2366.2	±	1.3	−2374	±	17	−2530	±	280
−	−	cst. / rot.	−2870.66	±	0.73	−2871.0	±	4.0	−2890	±	81
−	+	cst. / rot.	−3384.83	±	0.83	−3386.7	±	7.6	−3440	±	120
+	−	cst. / rot.	−3475.74	±	0.73	−3478.5	±	5.4	−3526	±	81
−	−	rot. / rot.	−4297.9	±	1.6	−4317	±	32	−4460	±	200
−	+	rot. / rot.	−4952.1	±	1.7	−4970	±	27	−5180	±	230
+	+	cst. / rot.	−5092.18	±	0.72	−5094.2	±	7.8	−5150	±	130
+	−	rot. / cst.	−19602.3	±	4.2	−19710	±	200	−20430	±	440

Table 3: Logarithms of the evidence ratios ($\Delta \ln \mathcal{Z} \equiv \ln (\mathcal{Z}/\mathcal{Z}_{max})$) for the twelve model variants resulting from phase and sign ambiguities in the two-trajectory Regge model for $K^+\Sigma^0$ photoproduction. The results are listed in order of decreasing probability for a prior width of 100.

G^v	G^t	K^+/K^{*+} phase	$\pi = U(0, \pm 100)$			$\pi = U(0, \pm 1000)$			$\pi = U(0, \pm 10000)$		
−	−	rot. / cst.	0			0			0		
+	−	rot. / cst.	−0.50	±	0.42	−0.42	±	0.63	−0.55	±	0.97
+	+	cst. / rot.	−0.67	±	0.40	−0.63	±	0.63	−0.53	±	0.98
+	+	rot. / rot.	−0.71	±	0.43	−0.65	±	0.62	−0.7	±	1.1
−	+	rot. / rot.	−0.79	±	0.44	−0.77	±	0.68	−0.68	±	0.97
−	−	rot. / rot.	−0.86	±	0.41	−0.78	±	0.67	−0.73	±	0.87
+	−	rot. / rot.	−0.88	±	0.47	−0.79	±	0.68	−0.9	±	1.1
−	+	cst. / rot.	−1.01	±	0.44	−1.03	±	0.62	−1.1	±	1.1
+	−	cst. / rot.	−1.22	±	0.45	−1.10	±	0.68	−1.2	±	1.1
−	+	rot. / cst.	−1.23	±	0.46	−1.11	±	0.72	−1.09	±	0.98
−	−	cst. / rot.	−1.71	±	0.43	−1.65	±	0.65	−1.71	±	0.96
+	+	rot. / cst.	−1.97	±	0.48	−1.88	±	0.69	−1.87	±	0.99

Furthermore, the comparison with Jeffreys' scale (Table 1) indicates that the $p(\gamma, K^+)\Lambda$ data exhibit decisive evidence for the model variant with a positive vector and a negative tensor coupling constant, and a rotating phase for both trajectories. Indeed, the difference in $\ln \mathcal{Z}$ with the second-best model is around 24, amply exceeding the value of 5 required for a decisive statement. This result resolves the sign and phase ambiguity for $K^+\Lambda$ photoproduction, which previously could not be achieved using high-energy data alone [4]. Moreover, the result from this Bayesian analysis is consistent with the previous analysis of this particular model, but it did not require an additional analysis with data from the resonance region (for which $E_{lab}^\gamma \lesssim 3$ GeV). In other words, the Nested Sampling method requires less experimental data to reach the same conclusion as the χ^2-analysis which used a much larger set of $K^+\Lambda$ photoproduction data.
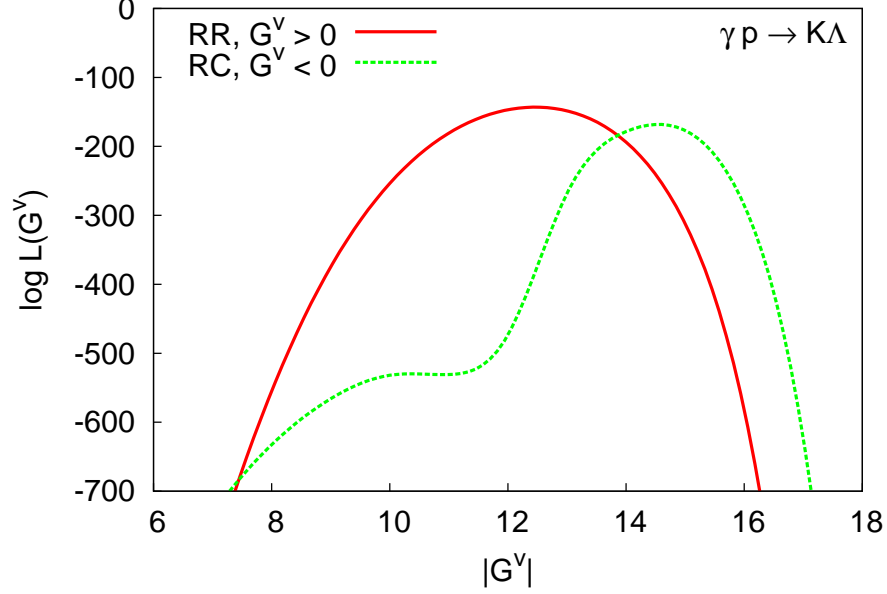
Figure 3: Comparison of the log-likelihood of $|G^v_{K^{*+}}|$ for the two best models. The model with the highest evidence is a rot./rot. (RR) model with a positive estimation value for this parameter. The second-best model, rot./cst. (RC), has a negative estimation value.

Apart from the relative probability of the different models, the Nested Sampling technique also provides us with an estimation value of the different parameters. In our best model, these values are:

$$\frac{g_{K^+\Lambda p}}{\sqrt{4\pi}} = -3.22 \pm 0.04$$
$$G^v_{K^{*+}} = 12.47 \pm 0.14$$
$$G^t_{K^{*+}} = -32.19 \pm 0.50. \tag{22}$$

Additional calculations demonstrate that the results do not change significantly when the $g_{K^+Y^0 p}$ coupling constant is allowed to deviate up to 40% from SU(3) predictions. For example, the values of $\Delta \ln \mathcal{Z}$ for the second and third $K^+\Lambda$ production models are -24.21 ± 0.73 and -77.09 ± 0.70 respectively, agreeing with the values found for 20% SU(3) symmetry breaking. The estimation values for the coupling constants are not affected. We conclude that the high-energy $p(\gamma, K^+)\Lambda$ data support a coupling constant $g_{K^+\Lambda p}$ compatible with a level of SU(3) symmetry breaking of at most 20%. The estimation value for $g_{K^+\Lambda p}$ deviates 15% from the SU(3) prediction.

Fig. 3 shows the log-likelihood of the parameter $G^v_{K^{*+}}$, integrated over the remaining two parameters, using a uniform prior between -100 and 100. The likelihood is determined by the high-energy $K^+\Lambda$ photoproduction data only.
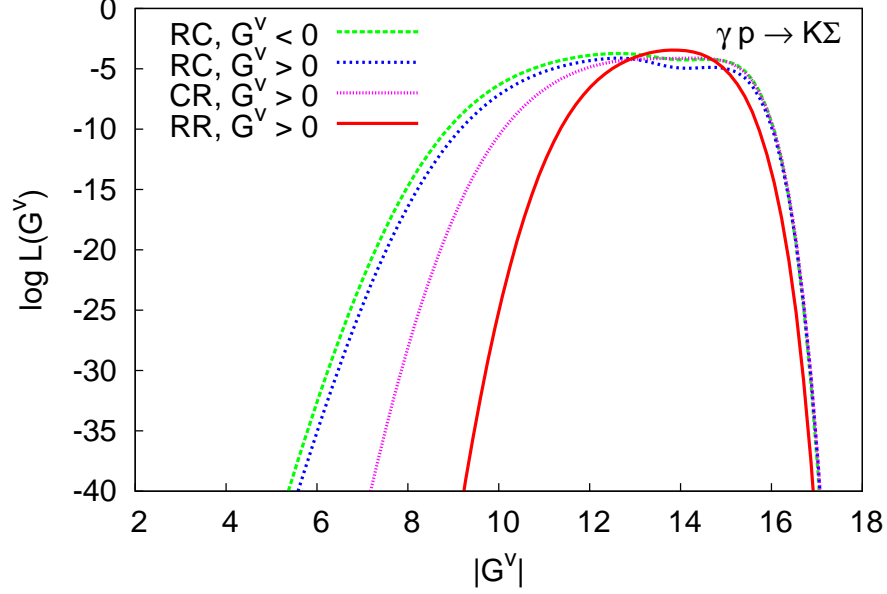
Figure 4: Comparison of the log-likelihood of $|G^t_{K^{*+}}|$ for the four best model variants. In contrast with Fig. 3, the model with the highest evidence value does not have the highest peak. Although the rot./cst. (RC) models surpass the cst./rot. (CR) and rot./rot. (RR) variants, the difference between the evidences is still too small to make a decisive statement.

The greater width and height of the peak of the rot./rot. model variant indicate that it should have a greater evidence than the alternative model.

The results for $K^+\Sigma^0$ production, listed in Table 3, are not as clear-cut as those for $p(\gamma, K^+)\Lambda$. The difference in $\ln\mathcal{Z}$ of the model variants is next to negligible. This was to be expected, as a smaller dataset provides fewer restraints on the models' parameters.

Even when the extensive set of resonance-region data is taken into account, the experimental data from the proton target alone does not allow us to single out one background model for the $K^+\Sigma^0$ channel [20]. However, a recent analysis of the corresponding reaction on the neutron, $n(\gamma, K^+)\Sigma^-$, was able to resolve the remaining ambiguity [27]. The models' parameters were converted from the proton to the neutron channel using isospin considerations.

The log-likelihood of the parameter $G^v_{K^{*+}}$ for $K^+\Sigma^0$ production is shown in Fig. 4. Note that despite boasting the maximal likelihood value, the rot./rot model has a lower evidence than the rot./cst. model.

5. Conclusions and outlook

Bayesian inference provides us with a promising tool for model comparison. We have demonstrated this by using the Nested Sampling algorithm to compute

the Bayesian evidence for different model variants of a Regge model for KY photoproduction. The results of this calculation indicate that there is decisive evidence for a $K^+\Lambda$ production model with $G^v_{K^{*+}} > 0$, $G^t_{K^{*+}} < 0$, and a rotating phase factor for both trajectories. This conclusion could not be drawn from the high-energy data set by means of the method of χ^2 minimisation.

For $K^+\Sigma^0$ production, the differences in evidence are too small to draw a decisive conclusion, and supplementary data is required to fully determine the background model for this channel.

The Nested Sampling method has many applications, both for the RPR model and for other research. One of these applications is the accurate estimation of model parameters as well as the elimination of nuisance parameters. More importantly, however, this method may provide us with a means to address the missing-resonance problem by calculating the probability of individual resonance contributions in a Bayesian framework. This is an approach we intend to explore in the near future.

Acknowledgements

References

[1] S. Capstick, W. Roberts, Prog. Part. Nucl. Phys. 45 (2000) S241–S331.

[2] M. Anselmino, E. Predazzi, S. Ekelin, S. Fredriksson, D. B. Lichtenberg, Rev. Mod. Phys. 65 (1993) 1199–1234.

[3] B. Saghai, J.-C. David, B. Julia-Diaz, T.-S. H. Lee, Eur. Phys. J. A31 (2007) 512–514.

[4] T. Corthals, J. Ryckebusch, T. Van Cauteren, Phys. Rev. C73 (2006) 045207.

[5] R. A. Adelseck, C. Bennhold, L. E. Wright, Phys. Rev. C32 (1985) 1681–1692.

[6] R. A. Adelseck, L. E. Wright, Phys. Rev. C38 (1988) 1965–1967.

[7] P. Mukherjee, D. Parkinson, A. Liddle, Astrophys. J. Lett. 638 (2) (2006) L51–L54.

[8] F. Feroz, M. P. Hobson, M. Bridges, Mon. Not. R. Astron. Soc. 398 (2009) 1601–1614.

[9] D. G. Ireland, et al., Phys. Rev. Lett. 100 (2008) 052001.

[10] M. R. Schindler, D. R. Phillips, Annals Phys. 324 (2009) 682–708.

[11] H. Jeffreys, Theory of Probability, 3rd Edition, Oxford University Press, Oxford, 1961.

[12] R. E. Kass, A. E. Raftery, J. Am. Stat. Assoc. 90 (430) (1995) 773–795.

[13] R. J. Barlow, Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences, Wiley, Chichester, 1989.

[14] D. S. Sivia, J. Skilling, Data Analysis – A Bayesian Tutorial, 2nd Edition, Oxford University Press, New York, 2006.

[15] J. Skilling, Bayesian Analysis 1 (4) (2006) 833–860.

[16] P. D. B. Collins, An Introduction to Regge Theory and High-Energy Physics, Cambridge University Press, Cambridge, 1977.

[17] M. Vanderhaeghen, M. Guidal, J. M. Laget, Phys. Rev. C57 (1998) 1454–1457.

[18] M. Guidal, J. M. Laget, M. Vanderhaeghen, Phys. Rev. C68 (2003) 058201.

[19] T. Corthals, D. G. Ireland, T. Van Cauteren, J. Ryckebusch, Phys. Rev. C75 (2007) 045204.

[20] T. Corthals, T. Van Cauteren, P. Vancraeyveld, J. Ryckebusch, D. G. Ireland, Phys. Lett. B656 (2007) 186–192.

[21] A. Boyarski, et al., Phys. Rev. Lett. 22 (1969) 1131–1133.

[22] D. J. Quinn, et al., Phys. Rev. D20 (1979) 1553–1583.

[23] G. Vogel, et al., Phys. Lett. B40 (1972) 513–516.

[24] R. A. Adelseck, B. Saghai, Phys. Rev. C42 (1990) 108–127.

[25] J. F. Donoghue, B. R. Holstein, Phys. Rev. D25 (1982) 2015–2018.

[26] M. Guidal, Ph.D. thesis, Université de Paris-Sud, U.F.R. Scientifique d'Orsay (1997).

[27] P. Vancraeyveld, L. De Cruz, J. Ryckebusch, T. Van Cauteren, Phys. Lett. B681 (2009) 428–433.